An overview of the V&V of Flight-Critical Systems effort at NASA

Author, co-author (Do NOT enter this information. It will be pulled from participant tab in MyTechZone)

Affiliation (Do NOT enter this information. It will be pulled from participant tab in MyTechZone)

ABSTRACT

As the US is getting ready for the Next Generation (NextGen) of Air Traffic System, there is a growing concern that the current techniques for verification and validation will not be adequate for the changes to come. The JPDO (in charge of implementing NextGen) has given NASA a mandate to address the problem and it resulted in the formulation of the V&V of Flight-Critical Systems effort. This research effort is divided into four themes: argument-based safety assurance, distributed systems, authority and autonomy, and, software intensive systems. This paper presents an overview of the technologies that will address the problem.

INTRODUCTION

These past few years, the US has started modernizing its air transportation system through an effort, known as the Next Generation Air Transportation System (NextGen). NextGen has been mandated by the US Congress with the goals [1] of retaining U.S. leadership in global aviation, expanding capacity, ensuring safety, and protecting the environment, through a cost beneficial implementation. To coordinate development and implementation of the Next Generation Air Transportation System by 2025, the US Congress created the Joint Planning and Development Office (JPDO), a multi-agency public/private initiative.

NextGen will require innovative changes to concepts of operation, demanding the implementation of complex software systems coupled with advanced hardware and communications capabilities. The validation and verification of these complex operational and technological developments has been identified by the JPDO has a critical gap for the implementation of NextGen [1]. Thus, The Research & Development activity R-1440, entitled "Applied Research on Complex Systems Validation and Verification", states this need as follows:

"*Applied research on the methods and algorithms to support the validation and verification of complex systems. Complex systems provide multiple functions that support many different operating models, environments and technologies and therefore require more advanced and integrated validation and verification methods and algorithms beyond those used for less complex systems. This research will support the development of complex systems, their risk assessment and eventual certification decisions."*

The JPDO also listed a corresponding enabler, called EN-3050 Advanced Complex System Validation and Verification Methods, which states the following:

"*Advanced tools and processes are developed to improve the verification and validation of complex systems and software. Improvements will focus on reducing the time and resources needed to conduct validation and verification as well as improving the quality of the results.*"

At the request of the JPDO, and as a member of the JPDO, the National Aeronautics and Space Administration (NASA) has been charged to address this critical gap in V&V of complex systems. Thus, the NASA Aeronautics' Aviation Safety Program has examined the research required to develop transformative safety validation and verification methods required to rigorously assure the safety of Next Generation Air Transportation System developments in a time- and cost-effective manner.

This article summarizes its findings and describes the resulting execution plan. We first motivate the need for better V&V for complex systems by describing how the increase in system complexity and the V&V cost drivers led us to divide the research into four important themes: argument-based safety assurance, distributed systems, autonomy & authority, and software intensive systems. The bulk of the paper is a description of each of these themes. It borrows heavily from the VVFCS study report, called Validation and Verification of Flight Critical Systems: Assessment of Critical Research Activities.

OVERVIEW

Current flight-critical systems are already so complex that they pose significant challenges to safety assurance; Table 1 shows the increase in software size (and complexity) for some well-known aerospace systems. For example, it is impossible to fully demonstrate safety in all possible operating conditions. Furthermore, the cost and time required by traditional V&V methods are prohibitive, and in many cases, they prevent the introduction of new technologies such as the use of adaptive algorithms in control. Yet, the operational improvements proposed under NextGen plans will increase the complexity of flight-critical systems and therefore reduce the current safety margins. Therefore, NextGen forces us to expand the traditional V&V strategy based on testing.

Table 1. Estimation of software complexity for some aerospace systems.

Mars Reconnaissance Orbiter	0.5 MLOC
Orion Primary Flight System	1.2 MLOC
F-22 Raptor	1.7 MLOC
Boeing 777	4 MLOC
Boeing 787	6 MLOC
F-35 JSF	>10 MLOC

The NASA study identified four general factors, which are outlined here:

- *The safety levels demanded of flight-critical systems are unique and unprecedented.*
- *Flight-critical systems span a range of design elements, including air, ground, and space. They involve human-human, human-automation, and automation-automation interactions with varying authority and autonomy constructs.*
- *Flight-critical systems are increasingly distributed (i.e., functions are delegated to individual components) and integrated (i.e., disparate components are providing information and functions to, and acting upon information and functions from, other components).*
- *Flight-critical systems are increasingly software-based, and that software is increasing in complexity.*

Besides the specific needs of complex systems, we are currently limited in our ability to perform V&V of flight-critical systems. For example, the current cost of V&V is enormous, especially given the requirements imposed by the certification process. Informal discussions with the FAA reveal that the verification cost to meet the DO-178B standards for level-A software (i.e., systems for which failures can have catastrophic consequences) can represent as much as seven times the development cost. Furthermore, V&V is not incorporated into the earliest stages of design, system development, during which V&V is actually less costly than in other stages, as shown in [2] by Barry Boehm and confirmed by industry during our study. Finally, V&V is often cited as an obstacle to innovation [3]. The cost and programmatic risk inherent to V&V can provide a business case against implementing new capabilities such as adaptive control, software with stochastic elements, or for novel air traffic concepts of operation that are highly distributed and interactive.

The VVFCS capabilities will advance safety assurance for flight-critical systems over their life cycle, to foster innovation within the air transportation system. The following four critical challenge areas have been identified:

- **Argument-based Safety Assurance**: ensuring safety in NextGen requires a comprehensive safety assurance framework applicable to NextGen operational concepts and to all its components and their interactions, wherever they reside - in the air, on the ground, or in space. Such a framework must address the increased complexity of NextGen and provide a consistent approach to safety assurance across all scales from specific components to airspace operational concepts.
- **Distributed Systems**: Critical aircraft system functions increasingly integrate (and are distributed across) multiple aircraft systems. Likewise, important airspace system functions increasingly integrate across multiple air and ground elements. This distributed nature can provide robustness, but can also be fragile when component interactions result in unintended consequences. Thus, capabilities to ensure the safety-critical properties of distributed systems need to be enhanced.
- **Authority and Autonomy**: tools are needed to assess whether advanced operational constructs properly assign authority and autonomy between automation and humans and whether such assignments are safe.
- **Software Intensive Systems**: The dependence on and complexity of software in the air and ground based components of the system will increase substantially with NextGen. There is a need to improve the effectiveness (and lower the cost) of assessing whether software meets safety objectives.

As mentioned above, new V&V methods should permit their application starting in the earliest stages of the design process and continuing throughout the remainder of the system's life cycle. This implies changes to V&V methods (e.g., quick, preliminary analysis methods early in design) and changes to design and life cycle management processes to integrally incorporate V&V.

ARGUMENT-BASED SAFETY ASSURANCE

The VVFCS study shows that ensuring safety in NextGen requires a comprehensive safety assurance framework applicable to NextGen operational concepts and to all its components and their interactions, wherever they reside - in the air, on the ground, or in space. Such a framework must address the increased complexity of NextGen and provide a consistent approach to safety assurance across all scales from specific components to airspace operational concepts.

Challenges especially relevant to the safety assurance of new technology and operational concepts include dissimilarities in approaches used for different types of systems, lack of uniform practice in requirements specification and validation, and lack of a systematic approach to incorporating lessons learned from operational experience about diverse factors including human performance, disturbances and degradations found in the operational environment, and unexpected interactions within the system and with other systems.

Different approaches and procedures are used for safety assurance of different types of systems, with notable distinctions between assurance methods for airborne and air traffic management systems [4]. Argument-based safety assurance methods used in other industries, often under the generic name of 'safety cases' [5,6], hold promise as a means of unifying and extending current safety assurance practices. A safety case for a system consists of explicit safety requirements, the evidence that these requirements have been met, and the argument linking the evidence to the requirements. Argument and evidence are both essential. An argument without adequate supporting evidence is unconvincing. A body of evidence without an argument is unexplained.

The research approach taken by VVFCS is to investigate whether argument-based assurance methods represent a practical solution for challenges facing safety assurance in civil aviation, and then to advance argument-based methods and tools to enable rigorous, and cost and time-effective safety assurance of future flight -critical systems. Novel elements of this approach include consideration of safety V&V throughout the entire system life cycle, from initial concept development, through design, implementation, modification, and operation, to final decommissioning. Another unique aspect of this research is that it will create a consistent single framework applicable to all the different types of systems envisioned for NextGen.

The three primary objectives are to:

- Provide a consistent means of safety assurance for flight-critical systems throughout their life cycle, i.e., concept development, implementation, operation, maintenance, modification and decommissioning;
- Enable improved comprehension of safety requirements, validation of those requirements, and verification of whether those requirements are satisfied, by the many participants in safety assurance, including designers, regulatory authorities, operators, and maintainers; and
- Investigate the requirements for evidence used in life cycle safety cases, including sources and types of evidence, characterizing and controlling for evidence quality, determining appropriate and inappropriate use and re-use of evidence, determining evidence limitations, and determining system and safety case requirements for support of life cycle safety case use.

DISTRIBUTED SYSTEMS

Over time, different philosophies have evolved concerning appropriate strategies for introducing and managing distributed systems in Aerospace systems. Different communication mechanisms may result in different failure manifestations and propagation. For each communication channel, an understanding of the various protection mechanisms that inhibit the potential for cascading failures is required. Additionally, there may be negative behaviors resulting from interference mechanisms through either unprotected, shared computational resources, e.g., shared memory or communications, or through coupling of physical elements, such as happens in a force fight between distributed controllers. Demonstration of the absence of the potential for adverse interactions is a significant technical challenge for the V&V of distributed systems.

Several architectural features may affect safe operation of a distributed system. These include the topology of the communication channels, the level of synchrony, and the capability to recover from disruptions and degradation within the system. Additionally, architectural choices impact the ability to maintain safety features as a system evolves throughout its life cycle.

Some recent incidents and accidents have resulted from unanticipated distributed system behavior in response to presumed rare subsystem failure modes and effects (eg, incidents in B777 and A330 ADIRUs). Therefore, VVFCS aims at providing better capabilities to model both the likelihood and system effects of disturbances and degradations that can adversely affect the behavior of flight-critical systems. There is a rich literature on theoretical foundations for distributed systems. However, there is evidence of a gap between these theoretical results and practical application [7,8]. For vehicle-specific systems, there is a need for advanced, reusable V&V capabilities to ensure safe management of both redundancy and shared resources in integrated, modular distributed systems. For airspace systems, there is a similar need for robust assurance for coordinated, distributed functional capabilities, e.g. separation assurance or merging and spacing.

The objective is to provide advanced analytical, architectural, and testing capabilities to enable sound assurance of safety-critical properties for distributed systems of systems. Drawing from several foundational research areas in formal methods, stochastic methods, test and evaluation methods, and architectural method, VVFCS is developing capabilities and supporting evidence to enable defensible, explicit assurance claims concerning safety implications introduced by technological advances in distributed systems, including the following:

- Validated models (both logical and stochastic) of failures, disturbances and degradation for current and anticipated aviation distributed systems
- Formal models of communication topologies, distributed algorithms (e.g. consensus, distributed resource management, mixed-criticality), and demonstration of resilience (through both analysis and test) to various combinations of disruptions, both permanent and transient, varying degrees of synchrony, fixed-topology (for vehicle systems) and dynamic topologies (for airspace systems)

- The development of approaches for modeling new systems decompositions and functional integrations that have been enabled by technological advances
- Tools and techniques to effectively transition abstract models of distributed algorithms/systems into practical engineering realizations.
- Techniques for assuring and preserving safety-critical properties of sub-systems within the context of higher-level systems, including allocating critical decisions to the correct locus of control.
- Capabilities for assuring functionally integrated distributed systems of systems from multiple vendors across a multi-decade system lifespan.

AUTONOMY & AUTHORITY

The US Air Traffic System (ATS), especially with future NextGen concepts of operation, is a complex system involving dynamic interactions among multiple actors that are largely governed through formal assignment of roles and responsibilities. These assignments of authority and autonomy are made at the design level, but are executed at the operational level according to each actor's view of their roles and responsibilities. Operationally, the system continuously adjusts for shortcomings in the assignment of authority and autonomy, for shortcomings in the capacity of actors to perform their assigned roles and responsibilities, and to optimize various performance factors such as capacity, environmental impact, and safety. This suggests that system safety should be derived not only from a predictable execution of assigned roles and responsibilities but also from checks and balances to ensure that the system operates as designed in the face of failures, disturbances and degradations. In this research area, the following definitions apply.

- ***Resilience***: The ability of a system to function safely in the presence of (compounding) disturbances.
- ***Authority***: The capability to act coupled with the responsibility to do so, within prescribed limits.
- ***Autonomy***: The ability to act and the freedom to do so without seeking permission, within prescribed limits.

The limits of authority and autonomy are typically based on the need for actors to clearly and unambiguously understand and predict each other's actions as well as their consequences. Who has what authority in what situation, and how the assignment of authority may affect safety, are important questions that need to be addressed when validating and verifying current and new airspace concepts of operations and flight-critical systems. The challenge is to assure, early-in-design, that authority and autonomy are clear, deadlock- and conflict-free, comprehensive, and consistent with agreed upon roles and responsibilities.

Given the flexibility of allocation of authority and autonomy in NextGen, it makes sense to expand the concept of safety from clear-cut, static, nominal conditions to the notion of resilience of a system to disturbances and degradations as enabled within communicating organizations. This is especially of concern when human flexibility is confronted with the rigidity of autonomous, non-human systems.

A surprising large number of commonly occurring problems can be found by rigorously considering safety early in design. V&V can be done on early designs via representations that unambiguously define safe critical configurations and that embody authority and autonomy requirements. Our challenge is to identify design and prototyping methods that show promise in how they represent authority and autonomy requirements and attributes and then extend state-of-art V&V to these methods. The extension of V&V methods must include their understanding of the critical safety attributes of authority and autonomy, including resiliency.

The main following research objectives are as follows:

- Safety analysis of models for the ATS using existing organizational modeling techniques
- Definition of formal representation of complex organizations that can support the V&V of Human-Machine Systems

- Develop techniques to analyze resilience of Authority and Autonomy in ATS
- Comprehensive toolkits for the formal analysis of organizations

SOFTWARE INTENSIVE SYSTEMS

The dependence on and complexity of software in the air and ground based components of the ATS will increase substantially with NextGen. We need to improve the effectiveness (and lower the cost) of assessing whether software meets safety objectives.

To date, the aviation industry has mostly relied on strong development process controls and traditional testing for the V&V of flight, ground, and air traffic control and management software. Strong process controls promote safety through detailed documentation of design decisions, coding, and V&V results; they may go as far as requiring repeatability, or even improvement, from one project to the next. In practice, process-oriented certification may have several disadvantages since it focuses more on documentation than demonstrating safety and eliminating potential problems. It also imposes a significant burden on the cost of developing software. The limitations of testing are well known; it is impossible to do exhaustive testing to find all bugs. Testing is also bad at finding bugs at the tail ends of fault distributions (e.g., bugs that show up only in specific, unusual flight conditions), even though these bugs may have a significant impact on safety.

VVFCS is focusing on the use of formal methods (model checking, static analysis, theorem proving, and more) because of their potential to guarantee a higher coverage of the software behavior than traditional testing or simulation. In the past, formal methods have been proven on small problems but have not scaled to the size of problem represented by current systems, especially given their distributed and heterogeneous nature; however, increases in processing power and recent advances in formal methods (such as compositional verification) indicate that formal methods may now be a valuable and tractable solution. Particularly for flight-critical systems, we will focus on numerical analysis, which is critical to the verification of flight software, the application of formal methods to model-based technology, and, compositional verification.

We also want to improve precision and coverage in testing and simulation. Improvements will come from formalizing the foundations of testing and simulation, investigating the use of mathematics for generating test artifacts, and combining them with other computer science techniques such as machine learning. For example, as pointed out in Owen, et al [9], statistical testing shows promise in increasing precision and coverage of testing and can be an alternative to formal methods, in particular for adaptive systems.

Research is also performed on studying if statistical testing can be used in conjunction with other techniques, e.g., formal methods, to increase software safety. For example, information gained through statistical testing can be used to guide the space in which formal methods techniques are applied. The complementary properties of theorem proving and model checking can be used in concert. A similar marriage can be made between formal methods and simulation-based statistical testing.

Main research elements for software intensive systems include the following:

- Research that improves analytical capabilities for building dependable software that makes significant use of numerical calculation.
- New capabilities that support reusable artifacts and publicly available libraries of definitions and theorems for invocation during formal analysis.
- Research to increase the precision and the coverage of testing while decreasing its cost.
- Research to enable the use of formal methods in argument-based safety assurance and in the context of model-based development methodologies widely adopted in industry.

SUMMARY/CONCLUSIONS

NASA has been asked to address a gap in the validation and verification of the complex systems needed to implement NextGen, the next generation of the Air Traffic System in the US. NASA has tasked the Aviation Safety program in the ARMD branch of studying how NASA can address this gap and it resulted in identifying the following four main research themes:

- **Argument-based Safety Assurance**: ensuring safety in NextGen requires a comprehensive safety assurance framework applicable to NextGen operational concepts and to all its components and their interactions, wherever they reside - in the air, on the ground, or in space. Such a framework must address the increased complexity of NextGen and provide a consistent approach to safety assurance across all scales from specific components to airspace operational concepts.
- **Distributed Systems**: critical aircraft system functions increasingly integrate (and are distributed across) multiple aircraft systems. Likewise, important airspace system functions increasingly integrate across multiple air and ground elements. This distributed nature can provide robustness, but can also be fragile when component interactions result in unintended consequences. Thus, capabilities to ensure the safety-critical properties of distributed systems need to be enhanced.
- **Authority and Autonomy**: tools are needed to assess whether advanced operational constructs properly assign authority and autonomy between automation and humans and whether such assignments are safe.
- **Software Intensive Systems**: the dependence on and complexity of software in the air and ground based components of the system will increase substantially with NextGen. There is a need to improve the effectiveness (and lower the cost) of assessing whether software meets safety objectives.

Despite budgetary constraints, the VVFCS research has started in FY'11 as an element in the System-wide Safety Assurance Technologies (SSAT) project in the Aviation Safety program. The goal of SSAT is to develop validated multidisciplinary tools and techniques to ensure system safety in NextGen and enable proactive management of safety risk through predictive methods.

REFERENCES

1. Next Generation Air Transportation System Integrated Plan, JPDO, December 2004
2. Boehm, B. 1981 *Software Engineering Economics*, as cited in DAA, 2008
3. System Transition: Dynamics of Change in the US Air Transportation System, Mozdzanowska, Aleksandra; Hansman, R. John; June 2008.
4. Hayhurst, Kelly J., Cheryl A. Dorsey, John C. Knight, Nancy G. Leveson, G. Frank McCormick, August 1999, Streamlining Software Aspects of Certification: Report on the SSAC Survey, NASA/TM-1999-209519.
5. Kelly, T. P. Arguing Safety - A Systematic Approach to Safety Case Management, PhD thesis Department of Computer Science, The University of York, United Kingdom, 1998.
6. Jackson, D., Thomas, M., Millett, L. I. (Eds). Software for Dependable Systems: Sufficient Evidence? National Research Council, Committee on Certifiably Dependable Software Systems (2007).
7. Chandra, T., Griesemer, R., Redstone, J. (2007) Paxos made live: an engineering perspective, *In Proceedings of the 26th ACM Symposium on Principles of Distributed Computing (PODC'07)*.
8. Driscoll, K., Hall, B., Sivencrona, H., Zumsteg. P. (2003) Byzantine Fault Tolerance, From Theory to Reality. Proc. 22nd International Conference on Computer Safety, Reliability and Security (SAFECOMP03), pp. 235-248, Edinburgh, Scotland, UK.
9. David Owen, Dejan Desovski, and Bojan Cukic. "Effectively Combining Verification Strategies: Understanding Different Assumptions." 17th International Symposium on Software Reliability Engineering (ISSRE '06), 10 pages, 2006.

CONTACT INFORMATION

Guillaume.P.Brat@nasa.gov

ACKNOWLEDGMENTS

This article has borrowed heavily (sometimes verbatim) from the report on the "Assessment of Critical Research Activities" for the "Validation and Verification for Flight-Critical Systems". Therefore, the author has generously borrowed from the prose of the following participants to the study: Sharon Graves, Steve Darr, Joseph Coughlan, Mike Shafto, Misty Davies, Kelly Hayhurst, Mike Holloway, Paul Miner, and other contributors such as John Ormes and Amy Pritchett (who was instrumental in starting and guiding this effort).

DEFINITIONS/ABBREVIATIONS

ADIRU	Air Data Inertial Reference Unit
ATS	Air Traffic System
FY	Fiscal Year
JPDO	Joint Planning and Development Office
NASA	National Aeronautics and Space Administration
NextGen	Next Generation Air Traffic Systems
SSAT	System-wide Safety Assurance Technologies project
V&V	Verification and Validation
VVFCS	Validation and Verification of Flight-Critical Systems